GEOPHYSICS, VOL. 58, NO. 4 (APRIL 1993); P. 553–563, 6 FIGS., 2 TABLES.

Quantitative determination of uncertainties in seismic refraction prospecting

Chuanhai Liu* and Joann M. Stock‡

ABSTRACT

We present a model of the propagation of refracted seismic waves in planar (horizontal or dipping) layered structures in which we quantify the errors from various sources. The model, called the (mixed) variance component model, separates the errors originating on the surface from those due to inhomogeneities of subsurface layers. The model starts with the assumption of homogeneous (constant-velocity) layers, but by taking the principal errors into account, variations from this model (including degree of velocity inhomogeneity, vertical velocity gradients, and gradational interfaces) can be identified.

A complete solution to the variance component model by Bayesian methods relies on the Gibbs sampler, a recently well-developed statistical technique. Using the Gibbs sampler and Monte Carlo methods, we can estimate the posterior distributions of any parameter of interest. Thus, in addition to estimating the various errors, we can obtain the velocity-versus-depth curve with its confidence intervals at any relevant point along the line.

We analyze data from a crustal-scale refraction line to illustrate both features of this method. The results indicate that the conventional linear regression model for the first arrivals is inappropriate for this data set. As might be expected, geophone spacing strongly affects our ability to resolve the heterogeneities. Differences in the amount of velocity heterogeneity in different layers can be resolved, and may be useful for lithologic characterization. For this crustal-scale problem, a velocity profile derived from this method is an improvement over simple linear interpretations, but it could be further refined by more comprehensive methods attempting to match later arrivals and wave amplitudes as well as first arrivals. The method could also be applied to smaller-scale refraction problems, such as determination of refraction statics, or constraints on the degree of probable lateral variations in velocity of shallow layers, for improved processing of reflection data.

INTRODUCTION

Presently, seismic refraction techniques are commonly used for two types of problems: determination of shallow structure, such as weathering thicknesses for statics corrections of reflection data (e.g., Palmer, 1986), and characterization of crustal-scale velocity structure (e.g., Mooney, 1989). Despite recent progress in data analysis in both of these areas, formal means of evaluating errors and uncertainties in forward modeling remain to be developed. This has become a major challenge in recent seismic work (Mooney, 1989). The quantitative measurements of the uncertainties, or random effects of the earth's interior, provide a good understanding of velocities of the corresponding inhomogeneous model. On the contrary, trying to investigate the properties, such as velocities, of the Earth's interior under the simplified assumption that the real medium is homogeneous may lead to serious errors (Červený and Firbas, 1984; Ojo and Mereu, 1986). To obtain uncertainties in the results of seismic refraction prospecting, the principal errors (or uncertainties) need to be classified. The paths of refracted waves lead to the classification of these errors into three types:

1) Surface error: Errors originating on the surface, including clock errors, station corrections and human error;

Manuscript received by the Editor January 21, 1992; revised manuscript received August 5, 1992.
*Department of Statistics, Harvard University, Cambridge, MA 02138.
‡Formerly Dept. of Earth and Planetary Sciences, Harvard University; presently at Seismological Laboratory 252-21, California Institute of Technology, Pasadena, CA 91125.

2) Layer-parallel error (*inhomogeneity*): Errors incurred when critically refracted waves (head waves) travel parallel to the interface between the layers;
3) Cross error (*inhomogeneity*): Errors incurred when the refracted rays travel through overlying layers down to the interface at which they are critically refracted, or from this interface up to the surface.

We treat these three types of errors as "random" errors, following Ojo and Mereu's (1986) arguments. The key point is that any "local fixed" effect on the traveltime of refracted waves in a relatively large formation can be viewed as a random effect on the total traveltime.

Some methods are available for numerical modeling of seismic body-wave fields in inhomogeneous anisotropic media. Červený and Firbas (1984) discussed these methods and provided a method of numerical modeling and inversion of traveltimes of seismic body waves in inhomogeneous anisotropic media. Ojo and Mereu (1986) discussed the magnitude of small-scale velocity inhomogeneities that could account for apparent linearity of segments of first-arrival traveltime curves from refraction data. However, as far as we know, the quantitative measurements of the inhomogeneities (or the estimation of the magnitude of the uncertainties) have not, as yet, been addressed.

In this paper, we separate the different errors using a variance component model, which is based upon some realistic assumptions about the planar (horizontal or dipping) interfaces in the subsurface. The magnitudes of the uncertainties are expressed in terms of variance components. The distribution of the uncertainties is obtained numerically by a recently well-developed statistical method called the Gibbs sampler (Geman and Geman, 1984; Gelfand and Smith, 1990; Gelfand et al., 1990). An example, using data from a reversed refraction profile from central California, is used to illustrate the method.

MODELING

In conventional seismic refraction work, the initial arrival of seismic energy is most important to the data analysis, although later arrivals as well as wave amplitude information may be used if conditions are favorable. The most common representation of refraction data is a plot of the recorded waveforms versus time T, aligned at a position x (the receiver distance from the shot point). For a subsurface consisting of discrete homogeneous layers, the first arrivals on this type of plot form linear segments, as shown in Figure 1. For planar interfaces, the slopes and intercept times (intersections of the lines with the T-axis) are used to determine the depths and velocities of layers in which the refracted waves traveled.

Theoretically, the first-arrival times of waves traveling in the kth layer form a linear segment if the top of the kth layer and each layer above the kth layer are homogeneous and $V_1 \le V_2 \le \ldots V_m$, where $k = 1, 2, \ldots, m$, m is the total number of layers traveled by the refracted waves and V_k is the velocity of the kth layer. For convenience in deriving the model for the kth segment, we temporarily drop the index k for layer number as well as the index for the profile number in the reversed profile. The paths of refracted waves are shown in Figure 1, where $x_1, x_2, \ldots, x_n$ are, respectively, the distances of n adjacent receivers from the shot point O, which record the refracted waves from the top of the kth layer, and $T_1, T_2, \ldots, T_n$ are the corresponding observed first-arrival times of the corresponding waves traveling along the path $O \rightarrow A^{(1)} \rightarrow B^{(1)}$, $O \rightarrow A^{(1)} \rightarrow A^{(2)} \rightarrow B^{(2)}, \ldots,$ $O \rightarrow A^{(1)} \rightarrow \ldots \rightarrow A^{(n)} \rightarrow B^{(n)}$, respectively.

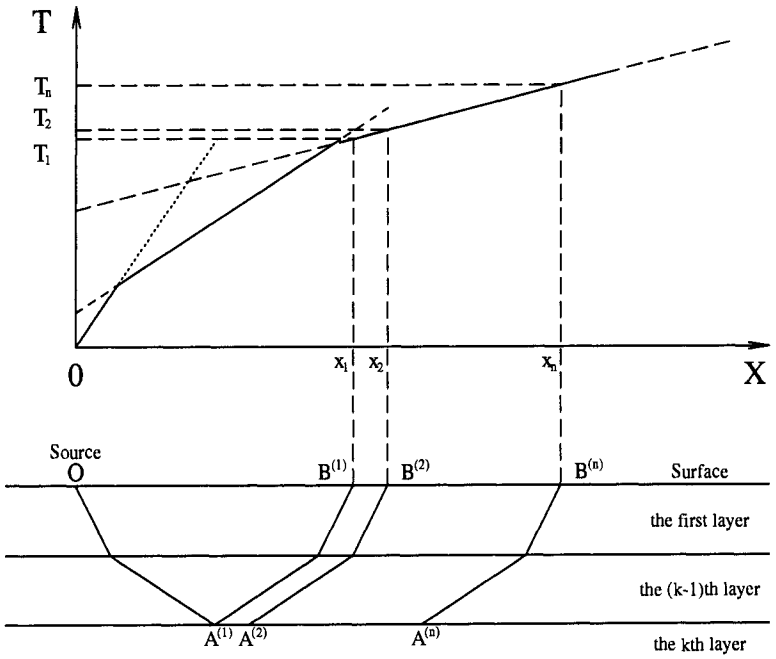


FIG. 1. Paths of refracted waves propagating in the earth's planar layered structures, illustrating the terminology used in this paper. Traveltime (T) is not reduced. Refracted arrivals along the path $O - A^{(n)} - B^{(n)}$ arrive at the receiver at $B^{(n)}$ at time T_n. Shotpoint for this profile is at O.

If we denote the slope of the kth linear segment to be μ and the corresponding intercept time as τ, then $\tau + \mu x_i$ is the expected traveltime of the first arrival at x_i. Let e_i denote the surface error, $e_{OA^{(i)}}$ and $e_{A^{(i)}B^{(i)}}$ the cross error, and $e_{A^{(i)}A^{(i)}}$ the layer-parallel error corresponding to x_i.

Thus, from Figure 1 we see that

$$T_1 = \tau + \mu x_1 + e_{OA(1)} + e_{A(1)B(1)} + e_1,$$

$$T_2 = \tau + \mu x_2 + e_{OA(1)} + e_{A(1)A(2)} + e_{A(2)B(2)} + e_2,$$

$$\cdots$$

$$T_n = \tau + \mu x_n + e_{OA(1)} + e_{A(1)A(2)} + \cdots + e_{A(n)B(n)} + e_n.$$

Our basic assumptions in this paper about the above errors in the refracted waves are the generalization of Tanar, Koehler, and Alhilali's assumption of "surface-consistency" (Taner et al., 1974). Specifically, we assume that:

1) Errors produced on the surface are "consistent," in other words, these errors are identically independently distributed if they are viewed as random variables. This assumption is quite realistic because we have no reason to know that the surface error at any geophone is correlated to that at any other geophone.
2) Inhomogeneity of any given interface is "consistent," in other words, the traveltime field at any small block in a given formation has the same possible deviation from the average traveltime field of the formation as that at any other small block that has the same length scale in the same formation. This assumption implies that the variance of layer-parallel error is proportional to the length of the top part of the kth layer along which the wave was refracted, because this error is the sum of "infinite" random errors along the kth top.
3) Also, if the dip of the structure from (b) is not large, the error from the source to the kth top and that from the kth top up to the surface can be assumed to be identically independently distributed.

Furthermore, a simple but realistic assumption about the distributions of the above errors is that these errors are normally distributed; the argument for this assumption is from the central limit theorem in probability theory (see the deterministic and indeterministic decomposition of Ojo and Mereu, 1986). Finally, we assume that the three different types of errors are independent; this is obviously an acceptable assumption.

Let $\mathbf{T} = (T_1, T_2, \ldots, T_n)'$, $\mathbf{X} = (x_1, x_2, \ldots, x_n)'$, σ_p = standard deviation of the layer-parallel error in one unit length (the distance unit of x), σ_c = standard deviation of the cross error and σ_s = standard deviation of the surface error. With the above assumptions, notations and algebraic derivation, $\mathbf{T}$ then is an n-dimensional normal distribution with mean $\mathbf{1}\tau + \mathbf{X}\mu$, and covariance $\sigma_c^2\underline{\mathbf{W}}_c + \sigma_p^2\underline{\mathbf{W}}_p + \sigma_s^2\underline{\mathbf{I}}$, i.e.,

$$\mathbf{T} \sim N(\mathbf{1}\tau + \underline{\mathbf{X}}\mu, \sigma_c^2\underline{\mathbf{W}}_c + \sigma_p^2\underline{\mathbf{W}}_p + \sigma_s^2\underline{\mathbf{I}}), \qquad (1)$$

where $\mathbf{1}$ is the n-dimensional vector with each component equal to 1, $\underline{\mathbf{I}}$ is the $n \times n$ unit matrix, and $\underline{\mathbf{W}}_c$ and $\underline{\mathbf{W}}_p$ are symmetric matrices. The lower triangular parts of $\underline{\mathbf{W}}_c$ and $\underline{\mathbf{W}}_p$ are respectively,

$$\underline{\mathbf{W}}_c = \begin{pmatrix} 2 & & & & \\ 1 & 2 & & & \\ 1 & 1 & 2 & & \\ \cdots & \cdots & \cdots & \cdots & \\ 1 & 1 & 1 & \cdots & 2 \end{pmatrix}$$

and

$$\underline{\mathbf{W}}_p = \begin{pmatrix} 0 & & & & \\ 0 & x_2 - x_1 & & & \\ 0 & x_2 - x_1 & x_3 - x_1 & & \\ \cdots & \cdots & \cdots & \cdots & \\ 0 & x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \end{pmatrix}.$$

For a reversed refraction profile with shot points denoted as 1 and 2, a distance R apart, there will be a linear segment on both profile 1 and profile 2 for each subsurface layer in the model. The notations for the parameters in equation (1) thus need to be specified with profile number and layer number, i.e., $\mathbf{T}$ is changed to $\mathbf{T}_{j,k}$, $\mathbf{1}$ to $\mathbf{1}_{j,k}$, $\mathbf{X}$ to $\mathbf{X}_{j,k}$, τ to $\tau_{j,k}$, μ to $\mu_{j,k}$, σ_c to $\sigma_{c,k}$, $\underline{\mathbf{W}}_c$ to $\underline{\mathbf{W}}_{c,j,k}$, σ_p to $\sigma_{p,k}$, $\underline{\mathbf{W}}_p$ to $\underline{\mathbf{W}}_{p,j,k}$, $\mathbf{I}$ to $\mathbf{I}_{j,k}$, and σ_s is kept the same [see assumption (1)], where $j = 1, 2$, $k = 1, 2, \ldots, m$. Note that $\sigma_{c,k}$ and $\sigma_{p,k}$ do not depend on the profile number because $\sigma_{c,k}$ and $\sigma_{p,k}$ describe the properties of the lithologic layers being sampled.

From equation (1) we obtain the following model for the reversed profile:

$$\mathbf{T}_{j,k} \sim N(\mathbf{1}_{j,k}\tau_{j,k} + \mathbf{X}_{j,k}\mu_{j,k}, \sigma_{c,k}^2\underline{\mathbf{W}}_{c,j,k} + \sigma_{p,k}^2\underline{\mathbf{W}}_{p,j,k} + \sigma_s^2\underline{\mathbf{I}}_{j,k}) \qquad (2)$$

where $j = 1, 2$ and $k = 1, 2, \ldots, m$. We add the restrictions: $\tau_{1,k} + R\mu_{1,k} = \tau_{2,k} + R\mu_{2,k}$, for $k = 1, 2, \ldots, m$ (equivalency of reciprocal times) and $\tau_{1,1} = \tau_{2,1} = 0$ (waves arriving from the first layer are direct arrivals).

The model described by equation (2) is called the (mixed) variance component model and the σ^2's are called variance components in statistics. If $\sigma_{c,k}^2 = \sigma_{p,k}^2 = 0$ for $k = 1, \ldots, m$, equation (2) describes the conventional linear model that can be solved by the least-squares method (Steinhart and Meyer, 1961).

ESTIMATION BY THE GIBBS SAMPLER

The estimation of the parameters in the mixed variance component model has been frequently addressed in the statistics literature (Hartley and Rao, 1967; Patterson and Thompson, 1971; Corbeil and Searle, 1976; Harville, 1977; Dempster et al., 1977). However, a complete solution to the problem [especially in the present case in which the quantities of interest, such as velocities, are complicated functions of the parameters in the model, equation (2)] relies on the Gibbs sampler (Geman and Geman, 1984, Gelfand and Smith, 1990, Gelfand et al., 1990)

The Gibbs sampler

The Gibbs sampler is an iterative method to generate random samples of the joint distribution $P(\Theta)$ of a random

vector $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_J)$ from its conditional probability distribution:

$$P(\Theta_i|\Theta_j, j \neq i)\ i = 1, \ldots, J, \tag{3}$$

where Θ_i, $i = 1, \ldots, J$, is a d_i—dimensional random vector and Θ is distributed over a $d = \sum_{i=1}^{J} d_i$—dimensional probability measure space Ω. In this particular case, Θ is a vector composed of the slopes, intercept times, sigma values, and missing data [see equation (8)] of interest to this problem.

The Gibbs sampler proceeds as follows. Given an arbitrary starting set of values $\Theta_1^{(0)}, \Theta_2^{(0)}, \ldots, \Theta_J^{(0)}$ with $P(\Theta^{(0)}) > 0$, we draw $\Theta_1^{(1)}$ from $P(\Theta_1|\Theta_2^{(0)}, \ldots, \Theta_J^{(0)})$, then $\Theta_2^{(1)}$ from $P(\Theta_2|\Theta_1^{(1)}, \Theta_3^{(0)}, \ldots, \Theta_J^{(0)}) \ldots$ and so on up to $\Theta_J^{(1)}$ from $P(\Theta_J|\Theta_1^{(1)}, \ldots, \Theta_{J-1}^{(1)})$ to complete one iteration of the scheme. After successive iterations, we would arrive at $(\Theta_1^{(n)}, \ldots, \Theta_J^{(n)})$. Under mild conditions, $\Theta^{(n)} \xrightarrow{d} \Theta \sim P(\Theta)$ as $n \to \infty$. Thus, for n large enough, say $n = N$, we can regard $\Theta^{(N)}$ as a simulated sample from $P(\Theta)$. Repetition of this sequence M times each to the Nth iteration generates M identically independently distributed samples $(\Theta_1^{(N,1)}, \Theta_2^{(N,2)}, \ldots, \Theta_J^{(N,M)})$. For a good review of the Gibbs sampler, please see Gelfand and Smith (1990). For the assessment of the convergence of the Gibbs sampler, please see Gelfand et al. (1990). More sophisticated methods of the assessment are presented in a series of technical reports in the Department of Statistics at Harvard University by Gelman and Rubin (1992), and Liu and Liu (1993).

With the generated samples, the density of Θ_i can be well estimated by the Rao-Blackwell type estimator

$$P(\Theta_i = \theta_i) = \frac{1}{M} \sum_{\ell=1}^{M} P(\Theta_i = \theta_i|\Theta_j = \Theta_j^{(N,\ell)}, j \neq i) \tag{4}$$

and random samples of a function $Y = f(\Theta)$ of Θ (e.g., the velocities, layer thicknesses, and dip angles as functions of the slopes and intercept times) can be created by

$$Y^{(N,\ell)} = f(\Theta^{(N,\ell)}),\ l = 1, 2, \ldots, M. \tag{5}$$

The inferences about $Y = f(\Theta)$ can thus be made based on the random samples from equation (5). For example, the confidence intervals of $Y = f(\Theta)$ can be constructed using the quantities of these samples; the estimates of its mean and variance are easy to derive

$$\bar{Y} = \frac{1}{M} \sum_{\ell=1}^{M} Y^{(N,\ell)} \tag{6}$$

and

$$Var(Y) = \frac{1}{M-1} \sum_{\ell=1}^{M} (Y^{(N,\ell)} - \bar{Y})^2, \tag{7}$$

respectively. In practice, we can aggregate the samples with $n > N$, say from $N + 1$ to $2N$, to obtain larger samples.

Conditional posterior probabilities of parameters in model, equation (2)

It is very difficult to obtain the joint posterior distribution of the parameters in model, equation (2). The full conditional posterior distributions can easily be obtained by introducing some "missing data": $\xi_{c,j,k} \sim N(0, \sigma_{c,k}^2 \underline{W}_{c,j,k})$, $\xi_{p,j,k} \sim N(0, \sigma_{p,k}^2 \underline{W}_{p,j,k})$ and $\xi_{s,j,k} \sim N(0, \sigma_s^2 \underline{I}_{j,k})$. The missing data ξ's correspond to the three types of errors. Model, equation (2), is therefore equivalent to the following model:

$$\mathbf{T}_{j,k} = \mathbf{1}_{j,k}\tau_{j,k} + \mathbf{X}_{j,k}\mu_{j,k} + \xi_{c,j,k} + \xi_{p,j,k} + \xi_{s,j,k}. \tag{8}$$

We use noninformative priors on the parameters τ's, μ's, and σ^2's (Box and Tiao, 1973); the full conditional posterior probabilities of τ's, μ's, σ^2's, ξ_c's, and ξ_p's are then:

1) $\tau_{1,1} = \tau_{2,1} = 0$, $\mu_{1,1} = \mu_{2,1}$ and $P(\mu_{1,1}|$ the rest) is a normal distribution with mean

$$(\mathbf{X}'_{1,1}\mathbf{X}_{1,1} + \mathbf{X}'_{2,1}\mathbf{X}_{2,1})^{-1}[\mathbf{X}'_{1,1}(\mathbf{T}_{1,1} - \xi_{p,1,1}) + \mathbf{X}_{2,1}(\mathbf{T}_{2,1} - \xi_{p,2,1})]$$

and variance

$$\sigma_s^2(\mathbf{X}'_{1,1}\mathbf{X}_{1,1} + \mathbf{X}'_{2,1}\mathbf{X}_{2,1})^{-1}.$$

For $2 \leq k \leq m$, $\tau_{2,k} = \tau_{1,k} + (\mu_{1,k} - \mu_{2,k})R$ and $P(\tau_{1,k}, \mu_{1,k}, \mu_{2,k}|$ the rest) is a normal distribution with mean

$$(\underline{\mathbf{X}}'_k\underline{\mathbf{X}}_k)^{-1}\underline{\mathbf{X}}'_k(\mathbf{T}_k - \xi_k)$$

and covariance

$$\sigma_s^2(\underline{\mathbf{X}}'_k\underline{\mathbf{X}}_k)^{-1},$$

where

$$\underline{\mathbf{X}}_k = \begin{pmatrix} \mathbf{1}_{1,k} & \mathbf{X}_{1,k} & \mathbf{0} \\ \mathbf{1}_{2,k} & R\mathbf{1}_{2,k} & \mathbf{X}_{2,k} - R\mathbf{1}_{2,k} \end{pmatrix},$$

$$\mathbf{T}_k = \begin{pmatrix} \mathbf{T}_{1,k} \\ \mathbf{T}_{2,k} \end{pmatrix},\ \xi_k = \begin{pmatrix} \xi_{c,1,k} + \xi_{p,1,k} \\ \xi_{c,2,k} + \xi_{p,2,k} \end{pmatrix}.$$

2) For $j = 1, 2$, $P(\xi_{c,j,k}|$ the rest) is a normal distribution with mean

$$\sigma_{c,k}^2(\sigma_{c,k}\underline{\mathbf{I}}_{j,k} + \sigma_s^2\underline{\mathbf{W}}_{c,j,k}^{-1})^{-1}(\mathbf{T}_{j,k} - \tau_{j,k}\mathbf{1}_{j,k} - \mathbf{X}_{j,k}\mu_{j,k} - \xi_{p,j,k})$$

and covariance

$$\sigma_s^2\sigma_{c,k}^2(\sigma_{c,k}^2\underline{\mathbf{I}}_{j,k} + \sigma_s^2\underline{\mathbf{W}}_{c,j,k}^{-1})^{-1},$$

and $\xi_{c,j,k}$ are independent for $k = 2, 3, \ldots, m$. $\xi_{c,j,1} = 0$ for $j = 1, 2$.

3) For $j = 1, 2$, $P(\xi_{p,j,k}|$ the rest) is a normal distribution with mean

$$\sigma_{p,k}^2(\sigma_{p,k}^2\underline{\mathbf{I}}_{j,k} + \sigma_s^2\underline{\mathbf{W}}_{p,j,k}^{-1})^{-1}(\mathbf{T}_{j,k} - \tau_{j,k}\mathbf{1}_{j,k} - \mathbf{X}_{j,k}\mu_{j,k} - \xi_{c,j,k})$$

and covariance

$$\sigma_s^2\sigma_{p,k}^2(\sigma_{p,k}^2\mathbf{I}_{j,k} + \sigma_s^2\mathbf{W}_{p,j,k}^{-1})^{-1}$$

and $\xi_{p,j,k}$ are independent for $k = 1, 2, \ldots, m$.

Note: the vectors and matrices above are in their "reduced form", i.e., the first components of the vectors and the first columns and rows of the matrices are omitted.

4) $P(\sigma_{c,k}^2|$ the rest$) = P(\nu s_{c,k}^2/\chi_\nu^2)$, and $\sigma_{c,k}^2$ are independent for $k = 2, \ldots, m$, where $\nu s_{c,k}^2 = \sum_{j=1}^{2} \xi'_{c,j,k}\mathbf{W}_{c,j,k}^{-1}\xi_{c,j,k}$, and $\nu = n_{1,k} + n_{2,k}$, where $n_{j,k}$ is the number of observations for the kth layer in the jth profile, and χ_ν^2 is chi-square distributed with ν degrees of freedom, i.e., $P(\sigma_{c,k}^2/\nu s_{c,k}^2|$ the rest$)$ is an inverted chi-square distribution (Box and Tiao, 1973) and $\sigma_{c,1}^2 = 0$ is a constant.

5) $P(\sigma_{p,k}^2|$ the rest$) = P(\nu s_{p,k}^2/\chi_\nu^2)$, and $\sigma_{p,k}^2$ are independent for $k = 1, \ldots, m$, where $\nu s_{p,k}^2 = \sum_{j=1}^{2}\xi'_{p,j,k}\mathbf{W}_{p,j,k}^{-1}\xi_{p,j,k}$, and $\nu = n_{1,k} + n_{2,k} - 2$, where $n_{j,k}$ is the number of observations for the kth layer in the jth profile, and χ_ν^2 is chi-square distributed with ν degrees of freedom [$\xi_{p,j,k}$ and $\mathbf{W}_{p,j,k}$ are in their reduced form; see (3) above].

6) $P(\sigma_s^2|$ the rest$) = P(\nu s_s^2/\chi_\nu^2)$, where $\nu s_s^2 =$

$$\sum_{j=1}^{2}\sum_{k=1}^{m}(\mathbf{T}_{j,k} - \tau_{j,k}\mathbf{1}_{j,k} - \mathbf{X}_{j,k}\mu_{j,k} - \xi_{c,j,k} - \xi_{p,j,k})'$$

$$\times(\mathbf{T}_{j,k} - \tau_{j,k}\mathbf{1}_{j,k} - \mathbf{X}_{j,k}\mu_{j,k} - \xi_{c,j,k} - \xi_{p,j,k})$$

and $\nu = \sum_{k=1}^{m} n_{j,k} - 2$. χ_ν^2 is chi-square distributed with ν degrees of freedom.

AN EXAMPLE

Data

We used the model, equation (2), to fit a data set recorded in the Great Valley, California, U.S.A. (Colburn and Mooney, 1986; Hwang and Mooney, 1986). The data set consists of a reversed refraction profile between shot point 12 and shot point 14 along the west side of the San Joaquin Valley. The two shot points are 96 km apart, i.e., $R = 96$ (km). First arrivals were picked by hand from "reduced traveltime" plots ($T - X/6.0$ versus X, where T stands for time and X is the source-receiver distance); the data points are shown in Figure 2. From the original reversed profile, we assume that the basic number of layers is 7. To evaluate the effect of segmentation of the data points, we compare the results from two alternative segmentations of the first arrival picks (Figures 2 and 3).

Estimation of the uncertainties

We applied the Gibbs sampler with $M = 100$ and set the starting points of the variance components to be the estimated variance of the conventional linear model and $\xi^{(0,\,l)} = 0.0$, $l = 1, \ldots, M$. One iteration of the Gibbs sampler was completed in the following order: τ's and μ's $\rightarrow \xi_{c,j,k}(j = 1, 2; k = 1, \ldots, 6) \rightarrow \xi_{p,j,k}(j = 1, 2; k = 1, \ldots, 6) \rightarrow \sigma_{p,k}^2$ $(k = 1, \ldots, 6) \rightarrow \sigma_{c,k}^2(k = 1, \ldots, 6) \rightarrow \sigma_s^2$. Using the convergence assessment of Gelfand et al. (1990) for σ^2's, we conservatively chose a value of $N = 100$. Continuing the Gibbs sampler up to $n = 2N = 200$, we obtained 100×100 random samples, which were used in the estimates of the parameters of interest. The estimates of the means and variances of uncertainties in model equation (8) by equations (6) and (7) are listed in Table 1a, where the numbers in parentheses are the corresponding standard variances.

To look at the different scale factors for the variance components in equation (2), let us go back to model (1). From model (1), we have

$$Var(T_i) = 2\sigma_c^2 + (x_i - x_1)\sigma_p^2 + \sigma_s^2. \qquad (9)$$

Therefore, we adjusted the uncertainty estimates according to their contributions to the total error in recorded traveltime (Table 1b). We see that apparently σ_s does not dominate the errors. Thus, an analysis based upon simple regression or least-squares fitting is inappropriate and may lead to a very different estimate of the uncertainties in the resulting curves of velocity versus depth for this data set (see the assumptions in Steinhart and Meyer, 1961).

We used equation (6) to compute the corresponding means, and the results of equation (5) to derive uncertainty limits on a profile of velocity versus depth, for a given location along the refraction line. The output from each sample of the Gibbs sampler was input into a standard dipping-layered-model refraction program, based on the equations of Mota (1954); means and uncertainty limits on the resulting output parameters (for instance, velocity as a function of depth; dip of individual layers) could then be obtained. We determined 95 percent confidence limits, rather than standard deviations, on velocity as a function of depth, because the probability distribution was often poorly approximated by a normal distribution, particularly near interfaces.

After the initial segmentation of the data points (Figure 2), we changed the assignment of points at the end of the layer 6 segment to the beginning of the layer 7 segment (Figure 3) to illustrate the effect this change in segment assignment would have on the results. Below, results from the two segmentation models (using the same first-arrival picks) are compared.

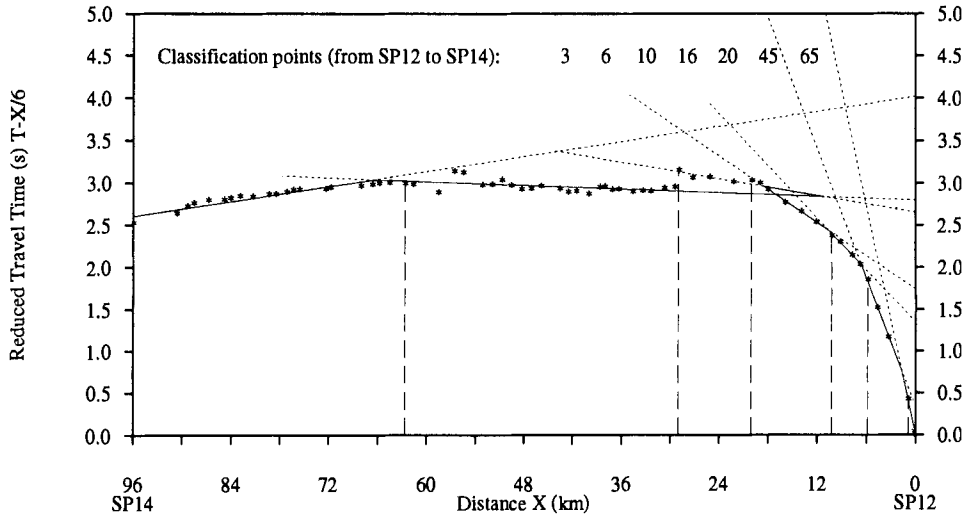
General results

The means and confidence limits on the resulting velocity–depth profiles (Figures 4 and 5) indicate with what confidence a layered structure can be resolved from these data. We have chosen to examine the resultant velocity profiles at two locations: 25 percent along the profile, or 24 km from each shot point. For each location, the uncertainties allow the velocity structure above 3 km depth to be fit equally well by a simple velocity increase with depth; this result is quite consistent with that of Ojo and Mereu (1986) who showed that apparently linear segments on traveltime curves might result from small random heterogeneities superposed on a linear velocity increase, particularly in the upper few km where crack closure might lead to a large velocity gradient. For the profile 24 km from SP 12 (Figure 4a), a layer of fairly constant velocity (4.31 km/s), which could have a slight velocity gradient, occurs from 3–6 km depth. The velocity profile from 6 to 7 km depth is relatively more poorly constrained, a layer with a velocity of 4.31 km/s may be interpreted to be present here, although the mean velocities

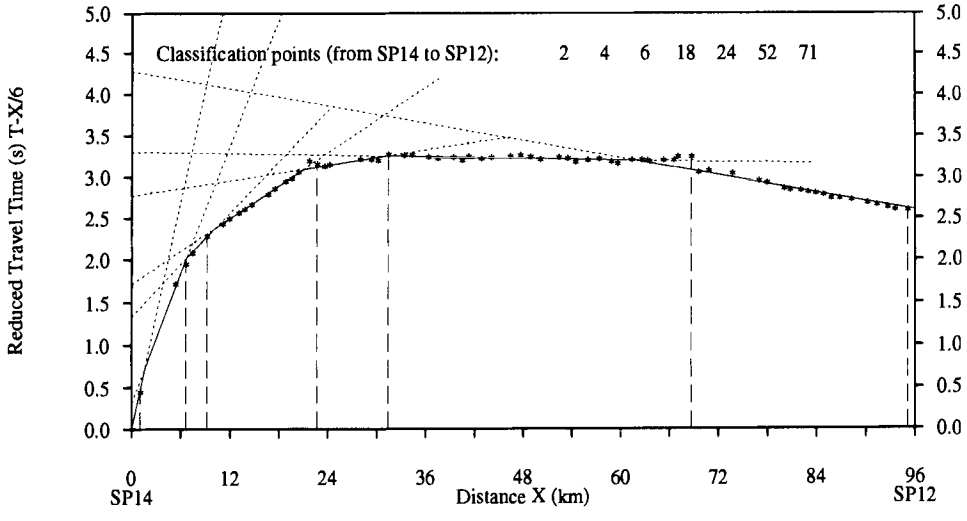
and the confidence limits allow this to correspond merely to a stronger vertical velocity gradient in this region. On the profile from shot point 14 (Figure 5), this layer is better defined due to its greater thickness. Velocities in the next two layers (6.01 km/s and 6.85 km/s, respectively), appear to be relatively constant, with the likelihood of only a small velocity gradient; the transition region between these two velocities, however, appears to have a much larger velocity gradient consistent with a gradational interface in this region. The uncertainty limits thus place some constraints on the magnitudes of allowable velocity gradients for given depths; as seen below, these results are in good agreement with best-fit curves from other more complicated methods of analysis. A cross-section with a layered-model interpretation (Figure 6) shows the "layers" labeled with their mean velocities but is somewhat misleading in terms of the velocity-depth profiles.

A comparison of the magnitudes of the errors determined for the data (Table 1) is useful, although we must first note that layers with fewer data points (layers 1, 2, 3, and to some extent, layer 5) have fewer degrees of freedom in these calculations and thus have less easily interpretable results. This can be seen from a comparison of the mean errors with their standard deviations (Table 1a); where the standard deviations exceed the size of the errors (Layers 1–5), conclusions may be difficult to draw. Because of the sizes of these mean errors and their standard deviations, the results from layers 4, 6, and 7 are most easily compared.

A comparison of the layer-parallel errors (normalized to path length in Table 1a) indicates the relative magnitudes of



(a)



(b)

FIG. 2. Picks of the first arrivals on the reversed profile between shotpoints 12 and 14, and initial assignments of data points to different segments. Data points are numbered going away from the shot point; the classification scheme indicated on the plot shows the observation number of the last point assigned to each segment. Vertical dashed lines down from each profile indicate the last data point assigned to each segment. The dotted linear segments were obtained by least-squares fitting. (a) Data from shot point 12 at the southern end of the profile. (b) Data from shot point 14 at the northern end of the profile.

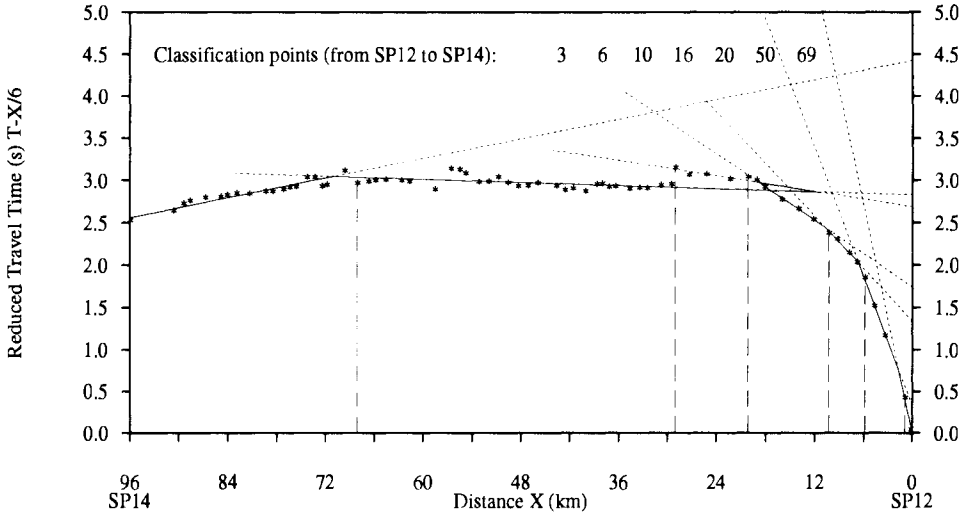
the errors associated with each layer; for example, layer 6 has twice the layer–parallel error of layers 5 or 7. This might be interpreted to suggest that layer 6 is more heterogeneous than the layers above and below it, which would be useful in distinguishing the lithology of layer 6 from layer 7. An alternative interpretation might be a structural one; for instance, the top of layer 6 is not as planar, on the scale of spacing of geophones, as the other interfaces. Colburn and Mooney (1986) attribute layers 6 and 7 to metamorphic and igneous rocks of the eastern foothills of the Sierra Nevada; if the layer–parallel errors are lithologically controlled, they suggest that layer 6 might be more compositionally variable (largely metamorphic rocks), whereas layer 7 might be less compositionally variable (largely igneous rocks). The layer–parallel errors in layer 4 (attributed to the Cretaceous Great Valley Sequence by Colburn and Mooney, 1986) are much lower in magnitude, possibly suggesting a more homogeneous composition for layer 4 than for the three layers below it.

The cross errors indicate the net error accumulated along the raypath above the layer in question; for the layers with a greater number of observations (layers 4–7) mean cross errors increase with depth to the layer, which is consistent with the increased path length.

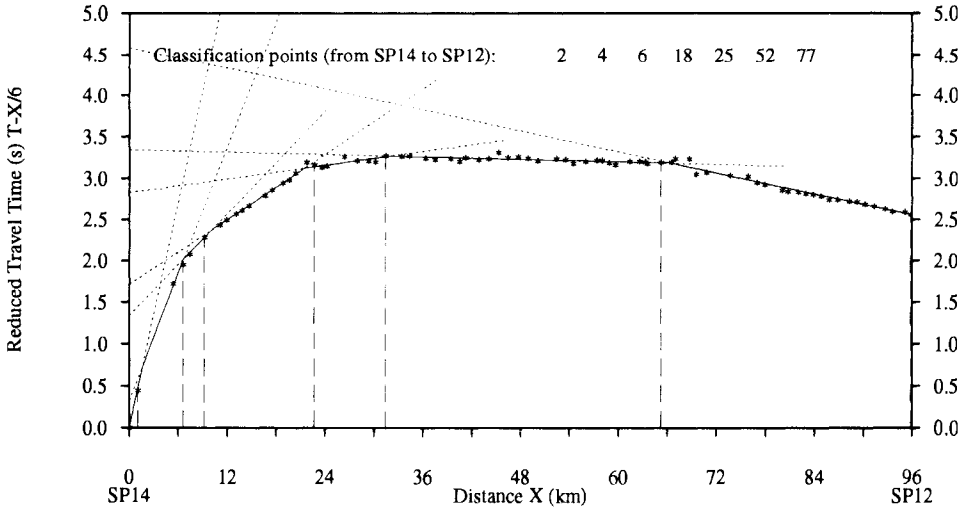
The total contributions of each type of error to the variance in the observed traveltime (Table 1b) shows that the contributions of the cross error and the layer-parallel error are about equal for each of layers 2–4, and subordinate to the surface error. For layers 5 through 7, however, the surface error is equal or less than the total contribution from the other error sources, which is reasonable given the longer path lengths for the observations from layers 5–7.

Comparison with the results of previous workers

The data set used in this example has been previously interpreted using forward modeling techniques that attempted to match many more details of the observed wave-



(a)



(b)

FIG. 3. Same data points and plotting conventions as in Figure 2, but with the data resegmented after least-squares fitting of lines to the data to achieve a better fit. The division of data points between segments 6 and 7 has been changed (compare to Figure 2).

Table 1(a). The estimated means and variances of the inhomogeneities from the Great Valley reversed profile. Units are as given for equation (1); for example, σ_p is normalized to unit path length.

Layer*	1	2	3	4	5	6	7
cross	—	0.0366	0.0056	0.0044	0.0116	0.0130	0.0223
$\sigma_{c,k}$		(0.0680)**	(0.0116)	(0.0055)	(0.0138)	(0.0094)	(0.0109)
layer-parallel	0.0212	0.0429	0.0052	0.0032	0.0121	0.0220	0.0109
$\sigma_{p,k}$	(0.0679)	(0.0703)	(0.0138)	(0.0041)	(0.0153)	(0.0075)	(0.0082)
surface error σ_s	0.0249	(0.0051)					

*Segmentation follows that of Figure 3.
**The number of parentheses is the corresponding standard deviation.

Table 1(b). The adjusted means of the uncertainties according to their contributions to total error of the recorded first arrivals.

Layer	1	2	3	4	5	6	7
adjusted cross error $\sigma_{c,k}$*	0.0000	0.0520	0.0079	0.0063	0.0165	0.0184	0.0315
adjusted parallel error $\sigma_{p,k}$**	0.0169	0.0550	0.0075	0.0074	0.0253	0.0893	0.0390
surface σ_s	0.0249						

*Adjusted by a factor of $2^{1/2}$ – see equation (9)
**The product of $\sigma_{p,k}$ and the average of $(x_i - x_1)^{1/2}$, $i = 2, \ldots, n$, in equation (1) in both profiles for each layer – see equation (9).

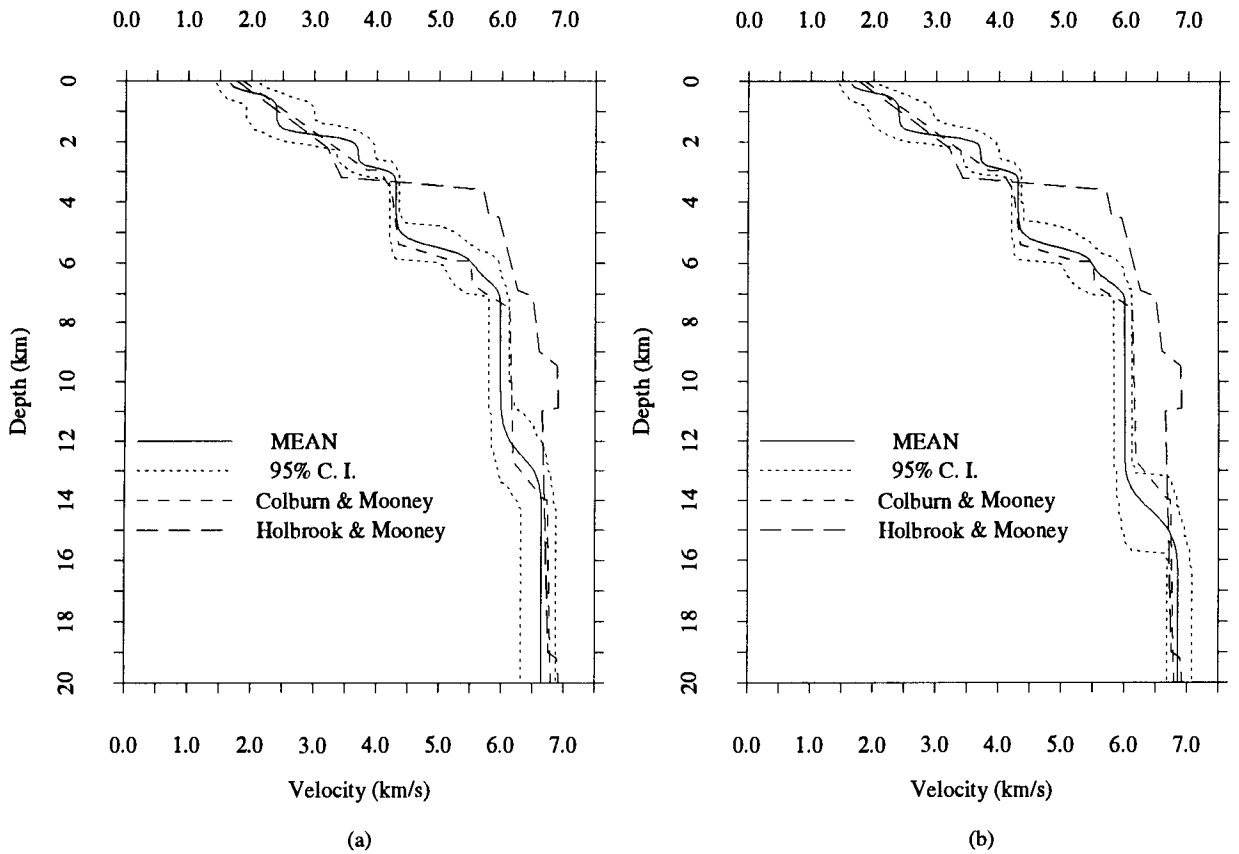


FIG. 4. Mean value of velocity versus depth (solid line) and 95 percent confidence limits (dotted lines) plotted for a position 24 km from shotpoint 12 and 72 km from shotpoint 14. (a) Input data are first arrival picks and segmentation shown in Figure 2. (b) Input data are the same first-arrival picks as in (a), but which the revised segmentation as shown in Figure 3. Velocity profile of Colburn and Mooney (1986) for the same position along this refraction profile, and Holbrook and Mooney (1987) for a parallel line to the east, are shown for comparison. Note that the 95 percent confidence regions in both cases encompass the results of Colburn and Mooney (1986) at depths above 8 km; for greater depths, Figure 4a appears to be a better fit to their profile.

forms than just the first arrivals used here. Colburn and Mooney (1986) used two-dimensional (2-D) ray tracing and attempted to match observed traveltimes of primary and secondary arrivals, as well as their amplitudes. They presented velocity profiles based on the same data that we used here. Hwang and Mooney (1986) used data from shot point 14 only, with a one-dimensional (1-D) technique, the modified reflectivity method, to determine a velocity profile and attenuation values consistent with the amplitudes and arrival times of primary and secondary phases and multiples. Holbrook and Mooney (1987) analyzed a parallel profile further east in the Great Valley that yielded a substantially different velocity structure than those of Colburn and Mooney (1986) or Hwang and Mooney (1986).

The confidence limits on our velocity profiles (Figures 4 and 5) allow us to compare our results with these previous results. We see that the profiles derived by Colburn and Mooney (1986) and Hwang and Mooney (1986) generally lie within the confidence regions of our results. We find relatively good agreement in terms of velocity gradients and gradational velocities at boundaries. Our mean velocity values cannot reproduce the existence of a low-velocity zone such as that postulated by Hwang and Mooney (1986), although the 95 percent confidence limits on our profiles would allow for the presence of a low velocity zone with the velocity structure they describe. The results of Holbrook and Mooney (1987) lie well outside the 95 degree confidence regions of our results from 2 km to about 12 km depth, suggesting that the confidence regions we determine are small enough to allow us to distinguish among various regional models.

Clearly, our technique cannot deal with many potential complications of the 2-D velocity structure beneath the refraction line; as mentioned above, low velocity zones would not be identified, nor would hidden layers or faults offsetting layers at depth. More detailed methods of analysis such as 2-D ray tracing, or comparison with synthetic seismograms, add an additional component of information that is critical to refining the velocity structure along the refraction profile. However, our method is a potentially useful "first-pass" approach. It is appealing because it requires knowledge of the first arrival times only, but yields information about the relative layer-parallel errors of the different layers, which may be related to lithology or interface geometry. It yields a velocity model (for dipping layers) with 95 percent confidence limits—extremely useful for a comparison between different velocity models. The resultant velocity profile is more realistic than a simple homogeneous plane layer model and could serve as a starting point for modification by more complete methods such as 2-D ray tracing or synthetic seismograms.

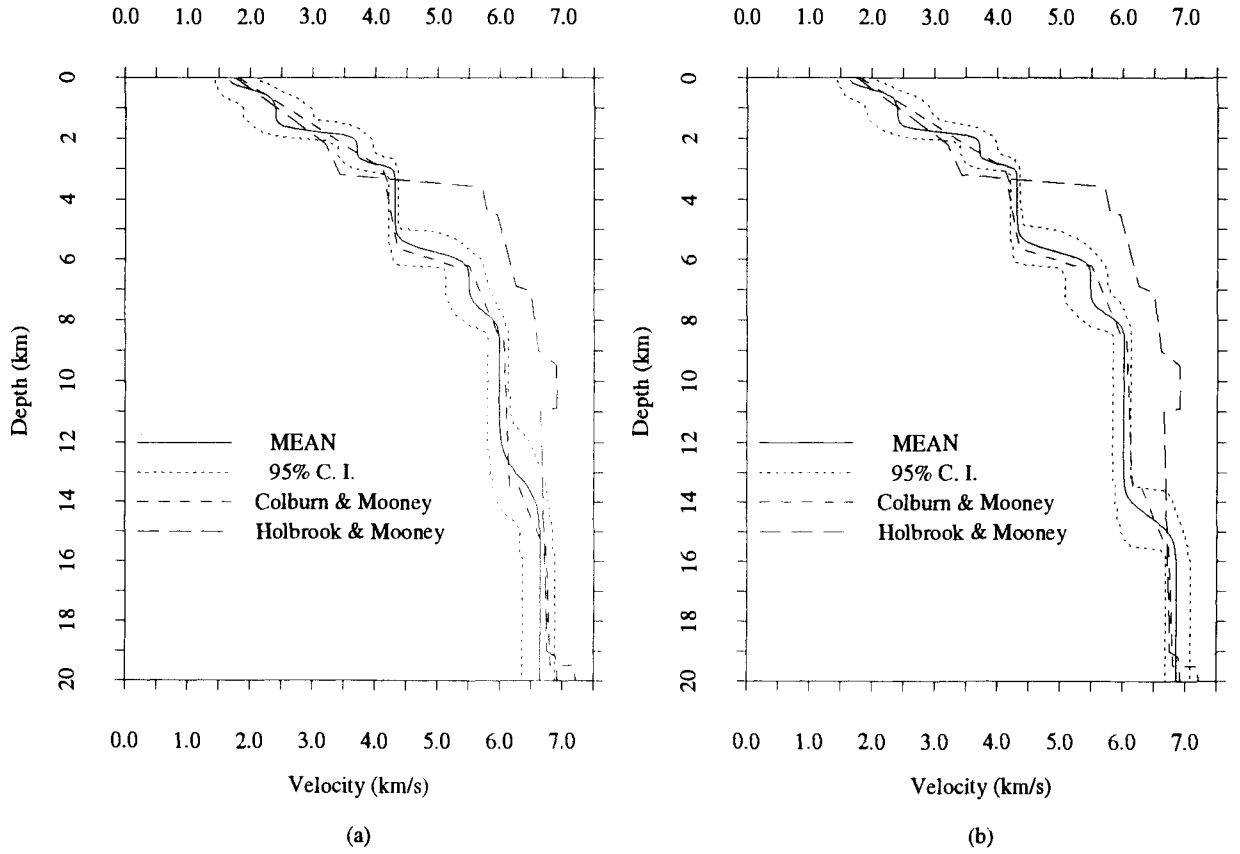


FIG. 5. Mean value of velocity versus depth (solid line) and 95 percent confidence limits (dotted lines) plotted for a position 24 km from shotpoint 14 and 72 km from shotpoint 12. (a) Input data are first-arrival picks and segmentation shown in Figure 2. (b) Input data are the same first-arrival picks as in (a), but with the revised segmentation as shown in Figure 3. For comparison, we show the velocity profiles of Colburn and Mooney (1986) for the same position along this refraction profile, Hwang and Mooney (1986) for an average of the profile from shotpoint 14, and Holbrook and Mooney (1987) for a parallel line to the east. Note that in Figure 5a the 95 percent confidence limits on mean velocity, derived here, encompass both Colburn and Mooney's (1986) and Hwang and Mooney's (1986) best-fit models down to 16 km depth. In Figure 5b the fit to Colburn and Mooney's (1986) profile is still within the uncertainties but Hwang and Mooney's (1986) profile is not.

Application to smaller-scale problems

The major use of refraction data in exploration is for providing absolute velocity and statics constraints for use with reflection processing, and we note here several ways in which this method might be relevant to such smaller-scale problems. We presume that in applying this method to reflection work, the geophone spacing would be more dense than for the crustal-scale problem shown here, and possibly various reversed profiles, with different endpoints, would be available along the same line. In this case, this method would constrain the range of allowable values of velocity to be used in determining static shifts [e.g., Yilmaz (1987), p. 227]. The confidence limits obtained on the velocity-depth curve would constrain the range of allowable velocity-depth functions derived from reflection processing, and place some constraints on the degree of velocity homogeneity expected or predicted for the various layers. A comparison of the cross-errors from the various layers would indicate which shallow layers are characterized by the most heterogeneity, and consequently would help to constrain the expected variation in amount of statics correction required along the line. The drawback of this method, when used with data from reflection sources, is that it requires the identification of the first breaks (first arrivals of the waveform). Difficulty in picking first breaks from some energy sources (e.g., vibroseis) may be reflected by an increase in the computed value of surface error.

SUMMARY

A realistic model of the propagation of refracted seismic waves in a planar-layered subsurface, based only on first arrivals from a reversed profile, has been constructed. The model has the following advantages. (1) It separates the three principal sources of errors and thus, provides quantitative measurements of the inhomogeneities in the subsurface; differences in the inhomogeneities of the different layers may indicate differences in lithology and/or interface geometry. (2) It takes the principal errors into account and therefore, inferences based on it are more reliable than those based on the simple conventional linear model in Steinhart and Meyer (1961). (3) By a Bayesian method, the estimation of the parameters of interest is easy to process using the Gibbs sampler and Monte Carlo methods. In the example, the estimates of the errors incurred during the layer-parallel and layer-crossing path of the raypath and the surface errors are presented, and confidence limits on velocity-depth curves are derived. Other parameters of interest, such as the thickness and dip angle of each layer considered in the model, could also be obtained using the equations in Mota (1954) and equations (5), (6) and (7).

ACKNOWLEDGMENTS

The data used in the example was originally provided by W. D. Mooney to J. Stock in 1984. C. Liu is grateful to his adviser Donald B. Rubin for his interest in the research, guidance, encouragement, and support from his NSF grant SES-88-05433. J. Stock was supported by NSF grant EAR-90-58217. We thank two reviewers and the Associate Editor for helpful comments.

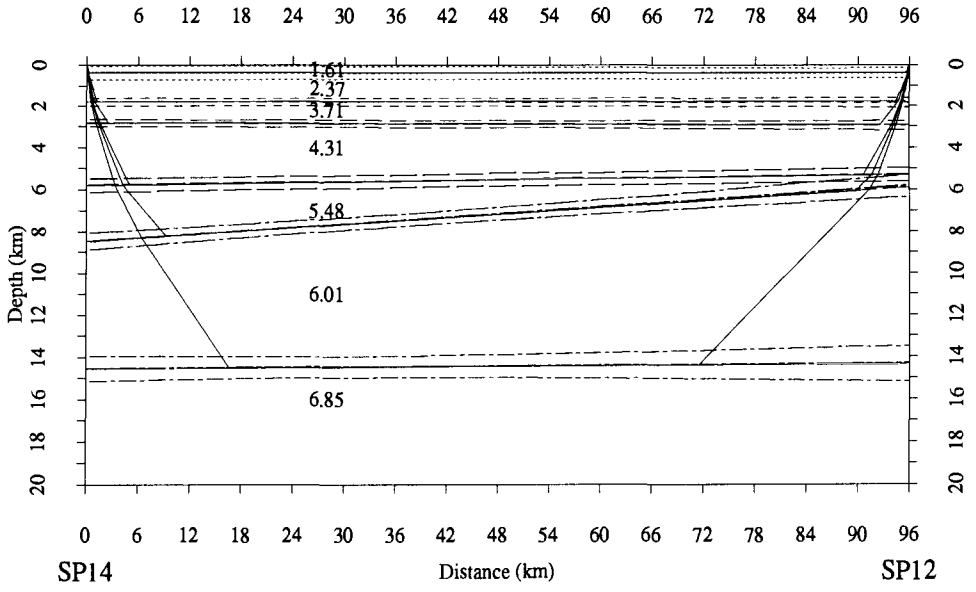


FIG. 6. Vertical cross-section along the refraction line, showing mean velocities and depths to interfaces derived from the revised segmentation of the observations (data shown in Figure 3). The 1-σ errors on the depths to the interfaces are shown, except for the lowest interface where both the 95 percent confidence limits and the 1-σ errors are shown. Note that the uncertainties on the depths to the deeper interfaces increase towards either end of the refraction profile (the 95 percent confidence limits or the 1-σ lines swing out from the mean line). Sample raypaths are shown from both shotpoints down to each interface; regions of the cross-section outside of these sample raypaths are unsampled and hence unconstrained. The apparent crossing of interfaces under shotpoint 12 is partly an artifact of this lack of sampling.

REFERENCES

Box, G. E. P., and Tiao, G. C., 1973, Bayesian inference in statistical analysis: Addison-Wesley Publ. Co.

Červený, V., and Firbas, P., 1984, Numerical modeling and inversion of traveltimes of seismic body waves in inhomogeneous anisotropic media: Geophys. J. Roy. Astr. Soc., **76**, 41–51.

Colburn, R. H., and Mooney, W. D., 1986, Two-dimensional velocity structure along the synclinal axis of the Great Valley, California: Bull. Seis. Soc. Am., **76**, no. 5, 1305–1322.

Corbeil, R. P., and Searle, S. R., 1976, Restricted maximum likelihood (REML) estimation of variance components in the mixed model: Technometrics, **18**, 31–38.

Dempster, A. P., Laird, N. M., and Rubin, D. B., 1977, Maximum likelihood from incomplete data via the EM algorithm (with discussion): J. Roy. Statis. Soc., Series B, **39**, 1–38.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. M., 1990, Illustration of Bayesian inference in normal data models using Gibbs sampling: J. Am. Statis. Assoc., **85**, 972–985.

Gelfand, A. E., and Smith, A. F. M., 1990, Sampling based approaches to calculating marginal densities: J. Am. Statis. Assoc., **85**, 398–409.

Gelman, A., and Rubin, D. B., 1992, Inference from iterative simulation using multiple sequences (with discussion): Statistical Science, **7**, no. 4, 457–511.

Geman, S., and Geman, D., 1984, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images: IEEE Trans. on Pattern Analysis and Machine Intelligence, **6**, 721–741.

Hartley, H. O., and Rao, J. N. K., 1967, Maximum likelihood estimation for the mixed analysis of variance model: Biometrika, **54**, 93–108.

Harville, D. A., 1977, Maximum likelihood approaches to variance component estimation and to related problems: Journal of the American Statistical Association, **72**, 320–340.

Holbrook, W. S., and Mooney, W. D., 1987, The crustal structure of the axis of the Great Valley, California, from seismic refraction measurements: Tectonophysics, **140**, 49–63.

Hwang, L., and Mooney, W. D., 1986, Velocity and Q structure of the Great Valley, California, based on synthetic seismogram modeling of seismic refraction data: Bull. Seis. Soc. Am., **76**, no. 4, 1053–1067.

Liu, C., and Liu, J., 1993, Comment on "Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods" by A. F. M. Smith and G. O. Roberts: J. Roy Statist. Soc., Ser. **B55**, no. 1, 82–83.

Mooney, W. D., 1989, Geophysical framework of the continental United States: Geol. Soc. Am. Memoir, **172**, 11–34.

Mota, L., 1954, Determination of dips and depths of geological layers by the seismic refraction method: Geophysics, **19**, 242–254.

Ojo, S. B., and Mereu, R. F., 1986, The effect of random velocity functions on the traveltimes and amplitudes of seismic waves: Geophys. J. Roy. Astr. Soc., **84**, no. 3, 607–618.

Palmer, D., 1986, Refraction seismics: Geophysical Press Ltd.

Patterson, H. D., and Thompson, R., 1971, Recovery of inter-block information when block sizes are unequal: Biometrika, **58**, 545–554.

Steinhart, J. S., and Meyer, R. P., 1961, Explosion studies of continental structure: Carnegie Institution of Washington Publ., **622**.

Taner, M. T., Koehler, F., and Alhilali, K. A., 1974, Estimation and correction of near-surface time anomalies: Geophysics, **39**, 441–463.

Yilmaz, O., 1987, Seismic data processing: Soc. Expl. Geophys.